



# Modelo predictivo basado en Machine Learning dirigido a PYMES de venta, caso de estudio Bluefields

*Predictive Modeling based on Machine Learning for retail SMEs, case study Bluefields*

Yosmaureen Naomi Villachica Pérez<sup>1</sup>  
Deyvon Kestner Ordoñez Cuthbert<sup>2</sup>  
Dexon Mckensy Sambola<sup>3</sup>

## Resumen

En la actualidad las pequeñas y medianas empresas (PYMES) se ven en el deber de hacer uso de la tecnología para afianzarse en el mercado, aprovechando las diferentes herramientas que surgen y que, por consiguiente, les permiten seguir siendo competitivas. Por ello, un modelo de predicción basado en Machine Learning constituye una herramienta primordial, su uso genera información que permite hacer toma de decisiones que afectarán de manera positiva los procesos de negocio. En este sentido, el objetivo principal del proyecto es desarrollar un modelo predictivo basado en Machine Learning dirigido a las PYMES de venta de la ciudad de Bluefields, Región Autónoma de la Costa Caribe Sur (RACCS), Nicaragua, que sirva como herramienta para maximizar las ganancias de estas empresas.

La metodología de desarrollo adoptada en la construcción del modelo de predicción fue “el modelo en cascada”, que mejor se adapta al contexto del proyecto. Esta metodología sugiere un enfoque sistemático y secuencial, disciplinado y basado en análisis, diseño, pruebas y mantenimiento.

Como resultado se obtuvo un modelo de predicción con los requerimientos necesarios para subsanar las deficiencias de estas empresas. El modelo está conformado por dos módulos, el primero es el SGPN's Companion que garantiza el almacenamiento, manipulación y seguridad de los datos, el segundo módulo es el Sistema Gestor de Procesos de Negocio (SGPN) o modelo de predicción, este permite realizar los estudios de procesos de negocio.

- 1 Ingeniera en Sistemas por la Bluefields Indian and Caribbean University. Correo: [nvillachica@gmail.com](mailto:nvillachica@gmail.com), ORCID <https://orcid.org/0000-0003-2411-6756>
- 2 Máster en Tecnologías de Análisis de Datos Masivos: Big Data por la Universidad de Murcia, España, responsable del área de Programación de la Bluefields Indian and Caribbean University. Correo: [deyvon.ordonez@bicu.edu.ni](mailto:deyvon.ordonez@bicu.edu.ni), ORCID <https://orcid.org/0000-0002-9109-5564>
- 3 Máster en Nueva Tecnologías en Informática, con Especialidad en Tecnologías Inteligentes y del Conocimiento con Aplicaciones en Medicina por la Universidad de Murcia, España, responsable de investigaciones de la Escuela de Informática y coordinador del grupo de investigaciones de IA & aplicaciones médicas de la misma escuela, Bluefields Indian & Caribbean University. Correo: [dexon.sambola@bicu.edu.ni](mailto:dexon.sambola@bicu.edu.ni), ORCID <https://orcid.org/0000-0002-3121-0831> mailto:

Recibido: 20/07/2021 - Aprobado: 24/02/2022

Villachica Pérez, Y. N., Ordoñez Cuthbert, D. K. ., & Mckensy Sambola, D. (2022). Modelo predictivo basado en Machine Learning dirigido a PYMES de venta, caso de estudio Bluefields. *Ciencia E Interculturalidad*, 30(01), 139–146. <https://doi.org/10.5377/rci.v30i01.14267>

**Palabras claves:** aprendizaje automático, Inteligencia de Negocios, algoritmo de aprendizaje, PYMES.

## Abstract

Nowadays, small and medium-sized enterprises (SMEs) are forced to make use of technology to gain a foothold in the market, taking advantage of the different tools that are emerging and that, consequently, allow them to remain competitive. Therefore, a predictive model based on Machine Learning is an essential tool, its use generates information that allows making decisions that will positively affect business processes. In this sense, the main objective of the project is to develop a predictive model based on Machine Learning aimed at sales SMEs in Bluefields city, Autonomous Region of the South Caribbean Coast, Nicaragua, which serves as a tool to maximize the profits of these companies.

The development methodology adopted in the construction of the predictive model was the “waterfall model”, which is the best suited to the context of the project. This methodology suggests a systematic and sequential approach, disciplined and based on analysis, design, testing and maintenance.

As a result, a predictive model was obtained with the necessary requirements to remedy the deficiencies of these companies. The model is composed of two modules, the first one is the SGPN’s Companion that guarantees the storage, manipulation and security of the data, the second module is the Business Process Management System (BPMS) or prediction model, which allows to carry out the business process studies.

**Keywords:** machine learning, Business Intelligence, learning algorithm, SMEs.

## I. Introducción

Tradicionalmente los sistemas de información se utilizan en las pequeñas y medianas empresas (PYMES) de ventas, para llevar el control y gestión de los productos, esto genera muchos datos que no son utilizados de la manera más apropiada. Una forma de aprovechar estos datos es mediante Machine Learning, una tecnología que permite crear modelos de aprendizaje para lograr entender mejor las preferencias del cliente. El objetivo principal del proyecto es desarrollar un modelo predictivo basado en Machine Learning dirigido a las PYMES de venta de la ciudad de Bluefields, que sirva como herramienta para maximizar las ganancias de estas empresas.

El proyecto se realizó debido a las dificultades observadas en cuanto al déficit de control y organización de las PYMES en la ciudad de Bluefields, por lo que ninguna de estas empresas tiene herramientas de esta naturaleza, cabe mencionar que para este estudio únicamente se tomó en cuenta las empresas que gestionan su proceso

mediante un sistema de información, debido a que el modelo hace predicciones en base a datos existentes para la toma de decisiones. Se llevó a cabo durante el 2020.

Entre las funcionalidades que presenta el modelo, están registrar y almacenar datos, brindar información de los procesos de negocio, realizar cálculos y predicciones, generar sugerencias de paquetes de productos e informes de la información procesada, de esta manera pretende garantizar una mejor organización y aprovechamiento de los diferentes recursos que posee el negocio. La mayor limitante fue el acceso de los datos de las empresas. Sin embargo, esta se superó con un conjunto de datos extraídos de base de datos externos. Cabe mencionar que estos cumplieron con las características necesarias para demostrar la funcionalidad del modelo construido.

## II. Revisión de literatura

Todas las ciencias están evolucionando debido al auge de las nuevas tecnologías de la informática, el mercadeo no es la excepción. Estamos en la era de la información, en donde muchos de los esfuerzos se centran en qué hacer con la cantidad excesiva de información generada, surgieron subcampos como la inteligencia de negocio que engloba un conjunto de tecnologías con el fin de mejorar la eficiencia de las empresas entre las cuales podemos mencionar el Big Data y Machine Learning.

El Machine Learning o aprendizaje automático es el subcampo de las ciencias de la computación y rama de la inteligencia artificial cuyo propósito radica en que las máquinas o computadoras aprendan automáticamente. Según Trujillo (2017), aprender en este sentido se refiere a la identificación de patrones complejos dentro de una gran cantidad de datos obtenidos mediante ejemplos, la experiencia o las instrucciones predefinidas.

La idea de aprendizaje se basa en un algoritmo que revisa los datos y es capaz de predecir comportamientos futuros adaptándose a la incorporación de información adicional y recalibrando los resultados. Por otro lado, la inteligencia de negocios es englobar un conjunto de conceptos, técnicas y herramientas con el fin de analizar y transformar los datos en información útil, que permite tomar decisiones estratégicas, tácticas y operativas más efectivas (Garcete et al., 2017).

En los últimos años estas tecnologías han tomado un rol importante en todos los sectores del comercio, en este momento nuestros esfuerzos se centran en la aplicabilidad de estas en las empresas de venta con el fin de rescatar las experiencias previas para aplicarlas en el contexto de interés de este proyecto.

Existen esfuerzos múltiples enfocados en demostrar la utilidad de estas tecnologías, como el estudio presentado por Garcete *et al.* (2017), que expone cómo la Inteligencia

de Negocios y Machine Learning ayudan de gran manera al pronóstico de la demanda con ayuda de algoritmos de aprendizaje.

Señala la importancia del uso de la tecnología para obtener beneficios y así minimizar riesgos e incertidumbre que puedan enfrentar las organizaciones dedicadas a las ventas. De acuerdo con los resultados experimentales se obtuvieron altas tasas de aciertos, haciendo pruebas exhaustivas con varios algoritmos de clasificación y evaluando con un método ampliamente aceptado. La técnica propuesta por los autores pretende que este nuevo modelo se convierta en una herramienta de apoyo en la toma de decisiones en el proceso de reposición de stock.

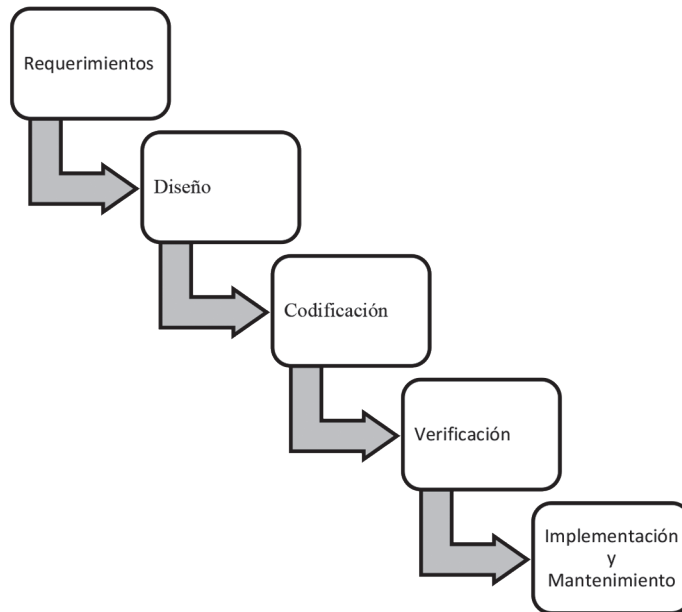
En un segundo trabajo presentado por Villacís Valarezo (2018), el autor presenta el desarrollo de una tienda virtual en la cual aplicó técnicas de Machine Learning para la automatización del proceso de promoción de productos de forma personalizada a cada consumidor. La metodología que se aplicó en el desarrollo del proyecto fue SCRUM. Utilizando Java y JavaScript específicamente las librerías React y Redux.

Por otro lado Kreplak (2018), presenta un modelo predictivo para pronosticar ventas futuras con razonable grado de exactitud con el fin de minimizar riesgos tales como sobreabastecimiento, que genera pérdidas, o la adquisición de una cantidad menor de productos de lo que se demanda, lo que produce un rápido agotamiento de productos y consumidores insatisfechos. Utilizó el algoritmo LGBM o Light Gradient Boosting Machine en Python. LGBM que es un algoritmo de tipo gradient boosting basado en árboles de decisión desarrollado en el marco de proyecto DMTK de Microsoft.

En el desarrollo de un modelo de predicción eficiente, es fundamental contar con tiempo suficiente para experimentar con todos los parámetros disponibles, procesar información y comentarios de otros usuarios que enfrentan el mismo problema predictivo. La revisión de la literatura permitió tener una postura más crítica en la selección de herramientas, tecnologías y metodología de desarrollo para la construcción del modelo propuesto.

### **III. Materiales y métodos**

Para la construcción del proyecto se adoptó la metodología de desarrollo en cascada, debido a que es la que mejor se adapta al contexto del trabajo. Esta sugiere un enfoque sistemático y secuencial, disciplinado y basado en análisis, diseño, pruebas y mantenimiento (Zumba y León, 2018). Las fases de esta metodología se aprecian en la figura 1.



**Figura 1. Fases de la metodología en cascada**

Fuente: elaboración propia

En la fase de requerimientos se hizo el análisis de las necesidades del cliente para determinar las características del modelo a desarrollar, se especifica todo lo que debe hacer el sistema sin entrar en detalles técnicos. Las herramientas de recopilación de datos utilizados fueron encuesta, entrevista y observación directa. En la fase de diseño se describió la estructura interna del modelo y las relaciones entre las entidades que lo componen, para ello se utilizó la herramienta de MySQL Workbench.

En la fase de codificación se programó los requisitos especificados haciendo uso de las estructuras de datos diseñadas en la fase de diseño. Al programar, también se realizó actividades de análisis de las condiciones, la creación de algoritmos y su implementación. Todo en base a lo determinado en la fase de requerimientos, para esto se usaron cuatro lenguajes principales: PHP, JS y HTML para el desarrollo del portal web y Python, para el modelo de predicción.

Después de tener el sistema terminado, en la fase de verificación se realizaron diferentes pruebas para asegurar que todas las funcionalidades estaban correctas. Para ello se usó un conjunto de datos extraídos de otra base de datos, mismos que cumplieron con todos los requerimientos para probar la funcionalidad del modelo de predicción.

El propósito de la fase de implementación y mantenimiento es instalar el modelo y mantener su valor a través del tiempo. Esto puede hacerse añadiendo nuevos requisitos, corrigiendo errores, mejorando la eficiencia o añadiendo nueva tecnología. El periodo de mantenimiento puede durar años. Debido a la naturaleza de este proyecto no se logró realizar esta fase. Sin embargo, en sustitución se ejecutó un conjunto de experimentos para validar el funcionamiento del modelo.

#### IV. Resultados y discusión

La situación actual de las PYMES de venta en la ciudad de Bluefields es bastante similar entre ellas, con variaciones leves. En el 50% de las tiendas que se estudiaron, un lapicero y un cuaderno son las herramientas utilizadas para el registro de ventas (no tienen procesos automatizados de ningún tipo).

Sólo la cantidad total de la venta es especificada y no se definen los productos vendidos en las transacciones. El porcentaje restante utiliza caja registradora (sistema automatizado) que sólo registra la cantidad a pagar. Sin embargo, ninguno tiene un modelo de predicción que apoye en la toma de decisiones.

El modelo resultante está conformado por dos módulos, SGPN's Companion y SGPN.

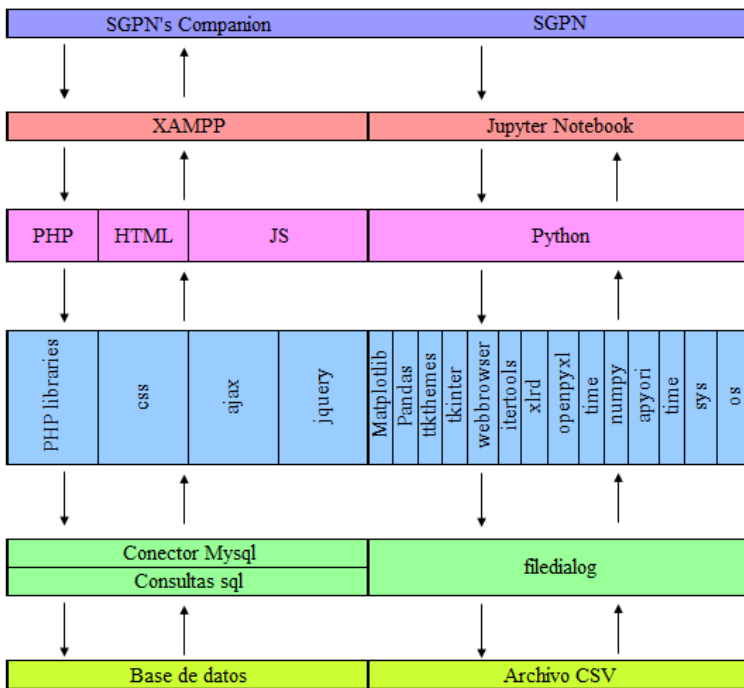
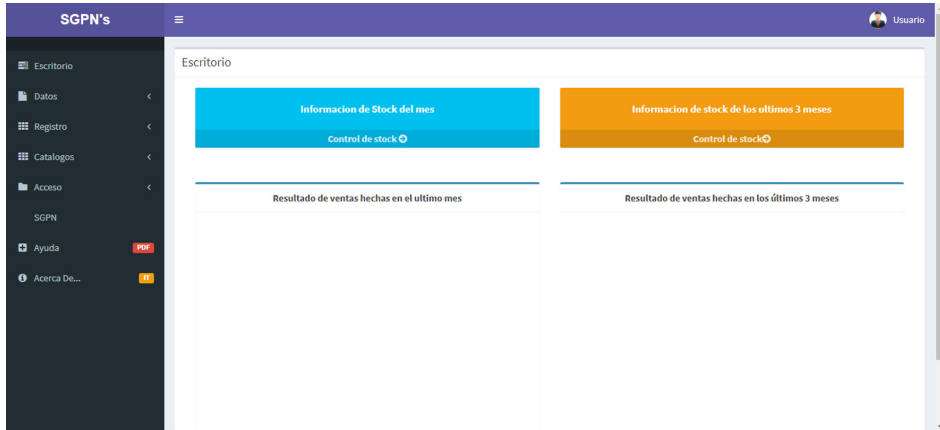


Figura 2: Arquitectura de modelo predictivo

Fuente: elaboración propia

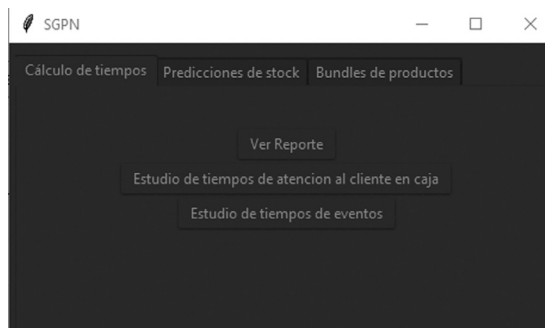
SGPN's Companion se encarga del almacenamiento, organización y gestión de los datos de los procesos de negocio de las empresas, esta se alojó en un servidor local. En ella se guarda información de los trabajadores, ventas y eventos realizados, registro de productos, adquisiciones de acciones y paquetes. A su vez, hay pestañas asignadas a diferentes tipos de información como ventas, eventos, trabajadores y productos. También permite editar las entradas y buscar la información, seleccionar los campos requeridos antes de guardarlo en un archivo con extensión CSV para su posterior análisis.



**Figura 3:** Vista principal del módulo SGPN's Companion

Fuente: elaboración propia

En la vista principal del módulo SGPN's Companion encontramos dos apartados, la revisión de acciones (stock) de los últimos 3 meses y el menú de navegación en la izquierda de manera vertical, en ella encontramos las pestañas de Datos, Registro, Catálogo, Acceso, SGPN, Ayuda y Acerca de SGPN o modelo de predicción, permite realizar los diferentes estudios de procesos de negocio. Los resultados de la predicción se reflejan en Jupyter Notebook, en ella se aprecian las gráficas, información y posibles recomendaciones, así como también las predicciones.



**Figura 4:** Vista principal del modelo predictivo

Fuente: elaboración propia

En la vista principal del modelo predictivo se encuentran los comandos para el estudio de tiempo de atención en la caja, cálculo de tiempo de eventos y ver reporte, el primer botón es para abrir archivos que el cliente haya registrado con anterioridad.

## V. Conclusiones

El modelo de predicción resultante es una herramienta que supone mejoras notorias en las PYMES de venta de la ciudad de Bluefields, los resultados de los experimentos realizados demostraron la funcionalidad del modelo propuesto. Con los módulos desarrollados se demostró que esta puede garantizar el alojamiento de los datos de la empresa, como la fiabilidad y confiabilidad de los datos que son la materia prima para realizar las predicciones que ayudarán a la toma de decisiones en las empresas.

Cabe mencionar que el sistema posee dos predictores para el análisis de dos procesos de negocio que son llevados a cabo por algoritmos de aprendizaje. El primero es el de predicciones de las acciones de valor (stock), el segundo es para predicciones de paquetes (bundles) de productos. Sin embargo, para garantizar el funcionamiento óptimo del modelo de predicción tendría que pasar por la fase de implementación y mantenimientos, donde esta sea alimentada de datos, en un tiempo más prolongado y en ambientes no controlados, sólo así se podría garantizar la depuración correcta y el funcionamiento óptimo del modelo. No obstante, los resultados experimentales fueron satisfactorios.

## VI. Lista de referencias

- Garcete, A., Benítez, R., Pinto-Roa, D., y Vazquez, A. (2017). *Técnica de pronóstico de la demanda basada en Business Intelligence y Machine Learning*. Simposio Argentino Sobre Tecnología y Sociedad, 193–202. [http://sedici.unlp.edu.ar/bitstream/handle/10915/64728/Documento\\_completo.pdf-PDFA.pdf?sequence=1](http://sedici.unlp.edu.ar/bitstream/handle/10915/64728/Documento_completo.pdf-PDFA.pdf?sequence=1)
- Kreplak, G. (2018). *Predicción de Ventas de Comestibles Corporación Favorita*.
- Trujillo Fernández, D. (2017). *Aplicación de metodologías Machine Learning en la gestión de riesgo de crédito*.
- Villacís Valarezo, G. A. (2018). *Desarrollo de Tienda Virtual que Proporcione Sugerencias de Compra por Medio de MACHINE LEARNING para la Empresa DEPORPAS S.A.* 103. <http://repositorio.ug.edu.ec/handle/redug/26702>
- Zumba, J., y León, C. (2018). Evolución de las Metodologías y Modelos utilizados en el Desarrollo de Software. *INNOVA Research Journal*, 3(10), 20–33.